MULTIMEDIA UNIVERSITY

STUDENT ID NO

MULTIMEDIA UNIVERSITY

FINAL EXAMINATION

TRIMESTER 1, 2015/2016

HPB2021 – BIOINFORMATICS PROGRAMMING II

(All sections / Groups)

05 OCT 2015
09-11 AM/PM
(2 Hours)

INSTRUCTIONS TO STUDENTS

1. This question paper consists of 7 pages including cover page.
2. Question one is **MANDATORY**. Attempt **THREE** out of **FOUR** remaining questions. All questions carry equal marks (10), and the distribution of the marks for each question is given.
3. Please write all your answers in the Answer Booklet provided.

QUESTION 1 – Mandatory [10 marks]

a) Write the output of the following codes. **[7 marks]**

i) [0.75 mark]

```
x = 3
if 2 > x :
   print 'First'
else :
   print  'Second'
   if 2 > x :
      print 'Third'
   print 'Fourth'
print 'Fifth'
```

ii) [1 mark]

```
words = 'this IS NoT EvEN'
print words.title()
print words.replace("IS", 'was')
print words.upper()
print words * 2
```

iii) [0.75 mark]

```
print  'left' "right" 'left'
```

iv) [1 mark]

```
fred = [ 'Alice', 'goes', 'to', 'market' ]
fred.extend([ 'with', 'Mike' ])
last = fred.pop()
fred.append('Fred')
print 'A:', fred
fred = [ 'On', 'Tuesday,' ] + fred
print 'B:', fred
fred.reverse()
print 'C:', fred
fred.sort()
print 'D:', fred
```

v) [1 mark]

```
for y in range(25, 20, -1):
   if y != 25:
         print y,
```

vi) [1 mark]

```
def fun(x, y):
   print x + y,
   print x * y,

fun(2, 3)
fun("2", 3)
fun("2", "3")
```

vii) [0.5 mark]

```
def what(n):
   if n == 0:
      return 0
   else:
      return n + what(n-1)
print what(5)
```

viii) [1 mark]

```
punctuations = '''!()-
[]{};:'"\,<>./?@#$%^&*_~'''

my_str = input("Enter a string: ")
new_str = ""
for char in my_str:
   if char not in punctuations:
      new_str = new_str + char
print new_str
>>>Enter a string: ("Hello!!!", he said ---&
he went.)
```

b) Provide R code for

i) Create a vector 'x' with integer values between 10-15.

ii) Create a matrix 'm_x' with 2 rows and 3 columns, enter the values into 'm_x' from vector 'x' by rows. (first 3 elements row 1, and next 3 elements row 2).

iii) Print column 3 of 'm_x' matrix. **[3 marks]**

Continue....

QUESTION 2 [10 marks]

a) Define a function "dig(-)" that receives a sequence and one restriction pattern and returns the positions where the enzyme cuts its pattern (Hint: the location where the restriction site starts in a sequence). You can return a list containing name of the enzyme and the position it occurs in a dna sequence.
e.g. EcoRI=['gaattc', 1,154,...] **[4 marks]**

b) DNA is a variable initiated to the following sequence: **[3 marks]**
DNA='gcatgacgttattacgactctgtcacgccgcggtgcgactgaggcgtggcgtcg'

i.Use ***replace()*** function to find the complement of dna sequence.
Hint: *complement_map = {'c': 'g', 'g': 'c', 'a': 't', 't': 'a'}*

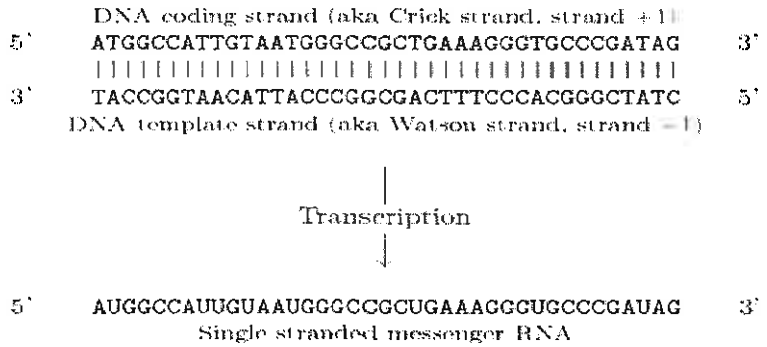
ii.Write the code to complement the dna sequence using ***translate()*** function.

c) Write the python code to open and read the file "***sequence.seq***" which is located in ***C:\files*** path on a Windows operating system. Search for those lines which start with C and print them out. Do not forget to eliminate extra next line at the end of each line before printing. **[3 marks]**

Continue....

QUESTION 3 [10 marks]

a) The diagram below shows the actual biological transcription process of the double- stranded DNA which encodes a short peptide.


DNA coding strand (aka Crick strand, strand +1)
5' ATGGCCATTGTAATGGGCCGCTGAAAGGGTGCCCGATAG 3'
3' TACCGGTAACATTACCCGGCGACTTTCCCACGGGCTATC 5'
DNA template strand (aka Watson strand, strand −1)
Transcription
5' AUGGCCAUUGUAAUGGGCCGCUGAAAGGGUGCCCGAUAG 3'
Single stranded messenger RNA

However, in Biopython, the program works directly with the coding strand. By this means, we can get the mRNA just by converting T -> U. Write a Biopython program to transcribe the DNA coding strand given below to an mRNA sequence.

DNA coding strand:
5'ATGGCCATTGTAATGGGCCGCTGAAAGGGTGCCCGATAG3'

Hint: use Bio.Seq & Bio.Alphabet **[5 marks]**

b) Write a program to translate the given DNA coding strand to Amino Acid : **5'**ATGGCCATTGTAATGGGCCGCTGAAAGGGTGCCCGATAG**3'**

[5 marks]

Continue...

QUESTION 4 [10 marks]

a) Write an R program to read a sequence from the user. The program counts the number of bases, "A","T","C", and prints the length of the sequence, bases, and the number of occurrences. **[4 marks]**

b) Write R code to import data from a comma delimited text file, '*mydata.csv*'. Assume the file has a header line. What is the data structure of the imported data? **[3 marks]**

c) Explain briefly what each of the following R functions does. **[3 marks]**

i) length(object)
ii) c(object, object, ...)
iii) ls()
iv) assign("even",c(2,4,6,8,10))
v) s3 = rep(a, 5)
vi) seq(4,10,by=2)

Continue...

QESTION 5 [10 marks]

a) The formulas for confidence intervals for the population mean depend on the sample size and are given below.

Confidence Intervals for μ	
For $n \geq 30$	For $n < 30$
$\overline{X} \pm Z\frac{s}{\sqrt{n}}$	$\overline{X} \pm t\frac{s}{\sqrt{n}}$
Use the Z table for the standard normal distribution.	Use the t table with df=n-1

Here we assume that the sample $\overline{X}$=5, s=2 and sample size=20. Use 95% confidence level to find the confidence interval. Write an R program to input the values and calculate the confidence interval **[6 marks]**

b) Calculate the following R vectors: **[3 marks]**

i) (1,2,3,....,19,20)
ii) (20,19,.....2,1)
iii) Create a vector of the values of $e^x\cos(x)$ *at x=3,3.1,3.2,....,6*
iv) ("label 1","label 2",.....,"label 30")

c) What commands can you use to get "***help***" in **R** programming? **[1 mark]**

End of Paper.